

Social Media Data in Active Mode research







Vincent Gong

EuroPride Amsterdam 2016

Jul 23 - Aug 7





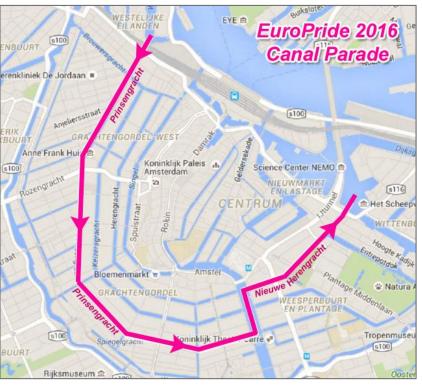




Canal parade Sat 6 August, 12:00 - 18:00

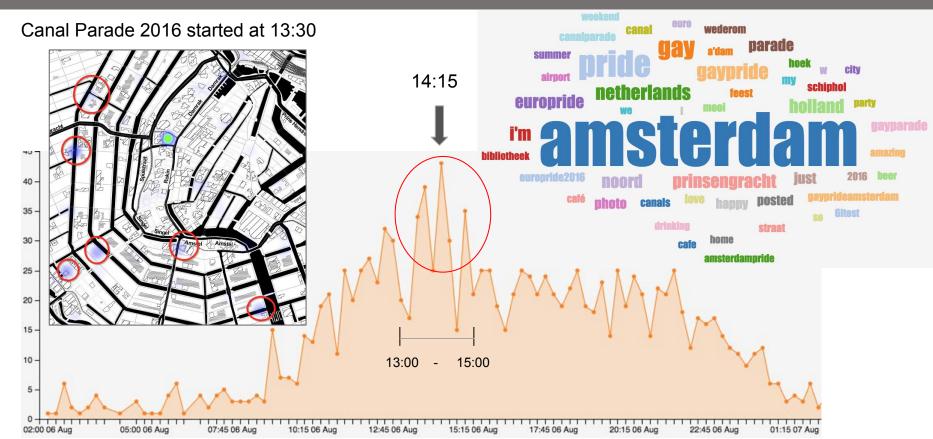






Canal parade Sat 6 August





Questions?



- What is social media data?
- Why to use it?
- To what extent can it be applied in Active Mode research? and how?

Agenda

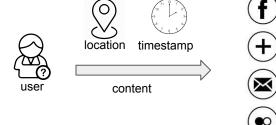


- What & Why social media data
- Information social media data provide
- Apply social media data in active mode research
- Societal Relevance

What is social media data



Social media, e.g.



- User profile
 - e.g. name, regi. date
- Content
 - e.g. text, image
- Timestamp
 - e.g. 10:00:00 Sep 6th, 2016
- Geo location
 - longitude, latitude

Why to use social media?



Urban Data

Census Records, Demographics, Spatial Statistics, Economic Data, Real-Estate Data, etc.

High data quality
High levels of accuracy,
completeness, and validity
Generally truthful
Semantic-by-design

Low refresh rate
Costly & laborious collection
methods
Non-scalable
Limited or no temporal
variability (static, semi-static)

Sensor & mobile Data

Physical Sensor Data, Mobile Phone Logs (CDRs), Transport Data, Energy Data, etc.

High levels of accuracy
High spatio-temporal resolution
High technology penetration
Generally truthful
Scalable & dynamic

Mostly proprietary
Very expensive to acquire
(CDRs)
Very expensive to deploy at the
city-scale
No semantics

Social media Data

Geo-localized Social Media Data from web platforms (e.g. Twitter, Instagram, 4SQ, Weibo, etc.)

High speed & refresh rate Created by people Enriched with annotations about places and human activities Scalable

Mismatch between the platform's scope and the application domain "Noisy" Biased (tech, social) Generally untrustworthy

Why to use social media?



Urban Data

Census Records, Demographics, Spatial Statistics, Economic Data, Real-Estate Data, etc.

High data quality
High levels of accuracy,
completeness, and validity
Generally truthful
Semantic-by-design

Sensor & mobile Data

Physical Sensor Data, Mobile

- Refresh rate,
- Capacity,
- Spatio-temporal resolution,
- Scalable,
- Enriched semantics

Ocherany truthiu

Scalable & dvnamic

Social media Data

Geo-localized Social Media Data from web platforms (e.g. Twitter, Instagram, 4SQ, Weibo, etc.)

New opportunities!

Enriched with annotations about places and human activities

Low refresh rate
Costly & laborious collection
methods
Non-scalable
Limited or no temporal



- Accuracy
- Biased
- Noisy

Mismatch between the platform's scope and the application domain "Noisy"
Biased (tech, social)
Generally untrustworthy

Agenda



- What & Why social media data
- Information social media data provide
- Apply social media data in active mode research
- Societal Relevance

Spatial—temporal feature EuroPride Aug 3rd to Aug 7th.

12:15 04 Aug

13:45 03 Aug

00:15 04 Aug

22:45 04 Aug





10:30 05 Aug

21:00 05 Aug

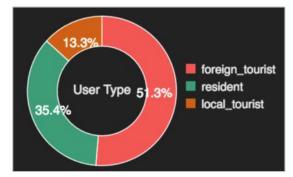
08:15 06 Aug

18:45 06 Aug

05:30 07 Aug

User type











Demographics

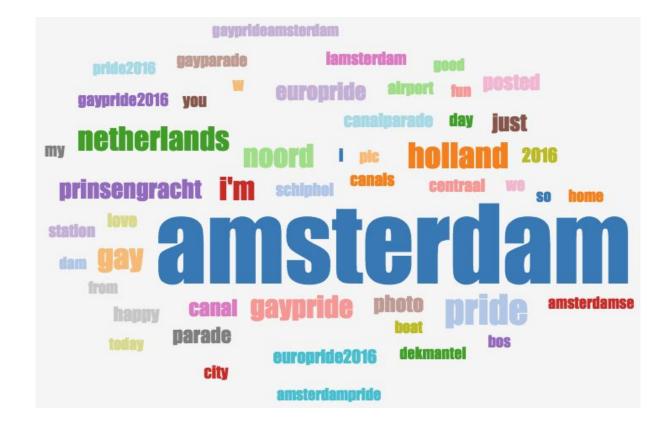




Content Analysis



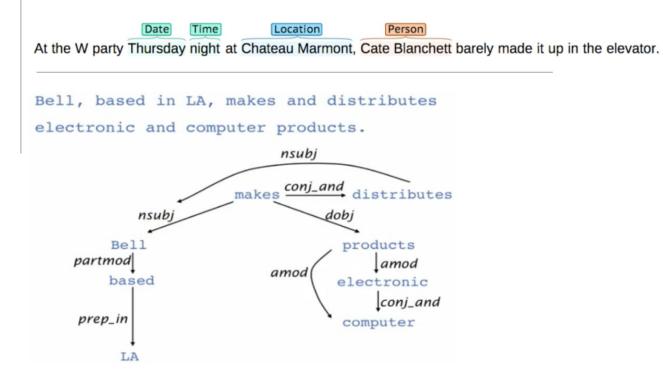
- Word occurrence
- Language analysis
- Image classification



Content Analysis



- Word occurrence
- Language analysis
- Image classification



Content Analysis



- Word occurrence
- Language analysis
- Image classification



Stage 0.90 Corn 0.03 Theater curtain 0.022 King penguin 0.012 Ear, spike 0.009



maillot 0.66 tank suit 0.10 two-piece 0.04 sarong 0.04 snorkel 0.02

Agenda



- What & Why social media data
- Information social media data provide
- Apply social media data in active mode research
- Societal Relevance

Social media data in active mode research



- Density, flow in a crowd event
 - sensor data vs social media data

- Activity scheduling
 - questionary data VS social media data

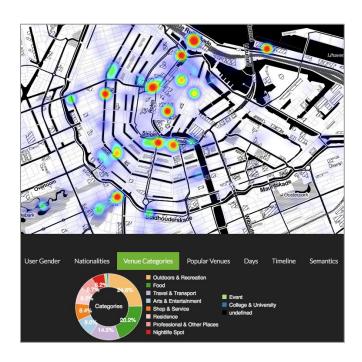
- City design and mgmt. w.r.t. pedestrians, cyclists
 - multiple data sources vs social media data



Societal Relevance



- Public administration
 - dashboard
 - data platform
- Retailer
 - promotion
 - adjust price strategy
 - organize events
- Individual
 - Recommendation, e.g. tripadvisor



Thank you and Q&A



